METHOD AND SYSTEM FOR GENERATING DATA PACKET IN DIFFERENT KINDS OF NETWORK

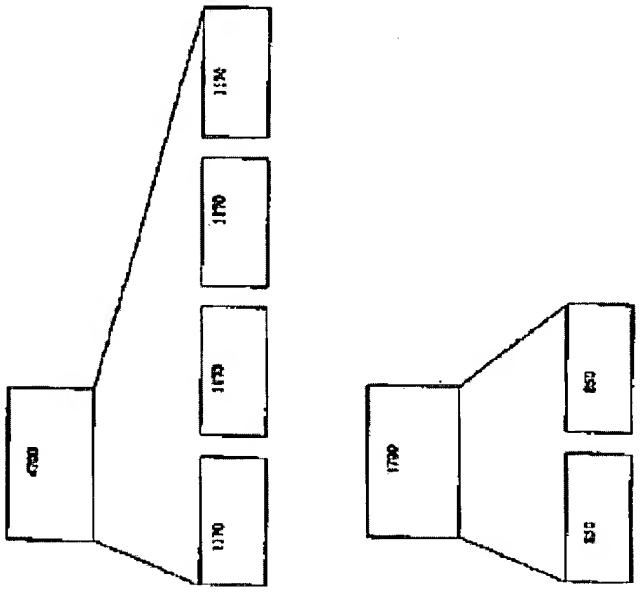
Patent number: JP11112574
Publication date: 1999-04-23
Inventor: MULLIGAN GEOFFREY
Applicant: SUN MICROSYST INC
Classification:
- international: H04L12/56; H04L12/46; H04L12/28; H04L29/06
- european:
Application number: JP19980175701 19980623
Priority number(s):

Also published as:
US6212190 (B1)
JP11112574 (A)

Abstract of **JP11112574**

PROBLEM TO BE SOLVED: To reduce quantity for fragmenting datagram into segments and to extend valid network band width associated with packet transmission on a network by comparing the size of a maximum transmission unit(MTU) that can be transmitted at a prescribed route with the sizes of transmitted packets and processing and transmitting the packets when the sizes of the packets are large.

SOLUTION: When the datagram segment is larger than MTU of a reception side network, a router fragments the datagram into plural smaller datagrams before the datagram is transmitted. The datagram of 4700 bytes is distributed to four datagrams provided with data of 1170 bytes and the datagram of 1700 bytes is divided into two smaller datagrams of 850 bytes for the network of 1500 bytes by MTU, for example. Data is uniformly distributed among the datagrams, the maximum fragmented size is reduced and the probability of fragmentation can be reduced.



(19)日本国特許庁（ＪＰ）　(12)**公開特許公報**（Ａ）　(11)特許出願公開番号

特開平11－112574

(43)公開日　平成11年(1999)４月23日

(51)Int.Cl.[6]	識別記号	ＦＩ	
Ｈ０４Ｌ　12/56		Ｈ０４Ｌ　11/20	１０２Ａ
12/46		11/00	３１０Ｃ
12/28		13/00	３０５Ｂ
29/06			

審査請求　未請求　請求項の数８　ＯＬ　（全 22 頁）

(21)出願番号　特願平10－175701

(22)出願日　平成10年(1998)６月23日

(31)優先権主張番号　０８／８８０２００

(32)優先日　1997年６月23日

(33)優先権主張国　米国（ＵＳ）

(71)出願人　591064003
サン・マイクロシステムズ・インコーポレーテッド
SUN MICROSYSTEMS, INCORPORATED
アメリカ合衆国　94303　カリフォルニア州・パロ　アルト・サン　アントニオ　ロード・901

(72)発明者　ジェフリー・マリガン
アメリカ合衆国・80920・コロラド州・コロラド　スプリングス・クローバーデイル　ドライブ・2175

(74)代理人　弁理士　山川　政樹

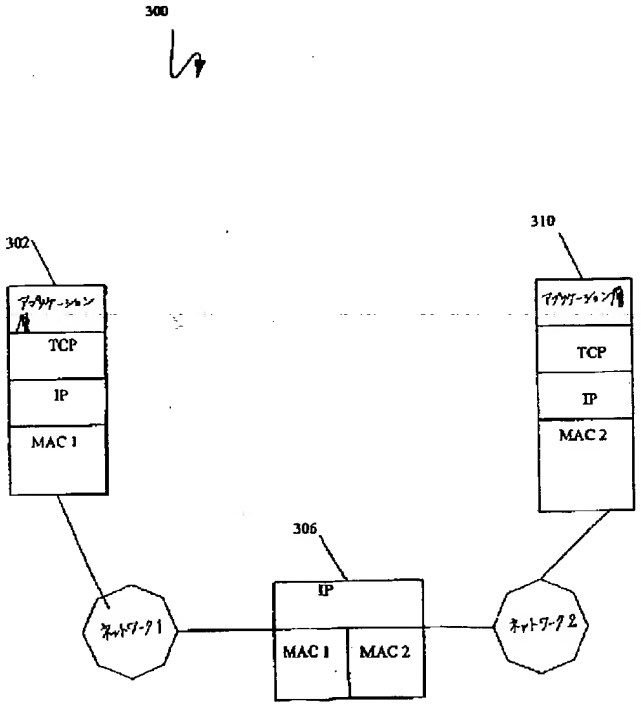
最終頁に続く

(54)【発明の名称】　異種ネットワークでデータ・パケットを生成する方法およびシステム

(57)【要約】

【課題】　ネットワーク上の異なるルートで送信されるパケットをより早く送ることができるように生成する改良された方法およびシステムを提供する。

【解決手段】　所定のルートで送信可能な最大送信単位（ＭＴＵ）を決定する。次に、ネットワーク上で送信する各パケットのサイズがＭＴＵのサイズと比較される。この比較でパケットがＭＴＵより大きいとわかると、パケットはさらに処理されてからルート上で送信される。その処理の特徴は従来のように最大ＭＴＵで分割するのではなく、個々のパケットのサイズ送ることができるサイズで、しかも均等のＭＴＵになるように分割する。



【特許請求の範囲】

【請求項１】　それぞれのルートが異なるサイズのパケットを伝送でき、パケットのサイズがパケット内に含まれる記憶単位の総数によって決まる、ネットワーク上の異なるルートで送信されるパケットのサイズを決定するためのコンピュータで実施する方法であって、
ネットワーク上で送信されるパケット内の記憶単位の総数をネットワーク上の所与のルートに関連する最大伝送単位（ＭＴＵ）と比較するステップと、
比較の結果がパケット内の記憶単位の総数がＭＴＵより大きいことを示している場合にパケットをさらに、
パケット内の記憶単位数をＭＴＵ記憶単位を保持できる１つまたは複数のパケット内に記憶された実質的に均等な単位のグループに分離することによって処理するステップとを含む方法。

【請求項２】　所定のルート上で送信可能な最大伝送単位（ＭＴＵ）を受信するステップをさらに含む請求項１に記載の方法。

【請求項３】　それぞれのルートが異なるサイズのパケットを伝送でき、パケットのサイズがパケット内に含まれる記憶単位の総数によって決まる、ネットワーク上の異なるルートで送信されるパケットのサイズを決定するように構成された装置であって、
ネットワーク上で送信されるパケット内の記憶単位の総数をネットワーク上の所与のルートに関連する最大伝送単位（ＭＴＵ）と比較するように構成された機構と、
比較の結果がパケット内の記憶単位の総数が所与のルートに関連するＭＴＵより大きいことを示している場合にパケットをさらに処理するように構成された機構と、
パケット内の記憶単位数をＭＴＵ記憶単位を保持できる１つまたは複数のパケット内に記憶された実質的に均等な単位のグループに分離するように構成された機構とを含む装置。

【請求項４】　所定のルート上で送信可能な最大伝送単位（ＭＴＵ）を受信するように構成された機構をさらに含む請求項３に記載の装置。

【請求項５】　それぞれのルートが異なるサイズのパケットを伝送できるネットワーク上の異なるルートで送信されるパケットを生成できるコンピュータ可読コードを記録した記録媒体であって、前記コードが、
ネットワーク上で送信されるパケット内の記憶単位の総数をネットワーク上の所与のルートに関連する最大伝送単位（ＭＴＵ）と比較するように構成された第１のコード部分と、
比較の結果がパケット内の記憶単位の総数がＭＴＵより大きいことを示している場合にパケットをさらに処理する第２のコード部分と、
パケット内の記憶単位数をＭＴＵ記憶単位を保持できる１つまたは複数のパケット内に記憶された実質的に均等な単位のグループに分離するように構成された第３のコード部分とを含むことを特徴とするコンピュータ・プログラムを記録した記録媒体。

【請求項６】　所定のルート上で送信可能な最大伝送単位（ＭＴＵ）を受信するように構成された第８のコード部分をさらに含む請求項５に記載の記録媒体。

【請求項７】　搬送波内で実施され、プロセッサによって実行されると、それぞれのルートが異なるサイズのパケットを伝送でき、パケットのサイズがパケット内に含まれる記憶単位の総数によって決まる、ネットワーク上の異なるルートで送信されるパケットのサイズを、
ネットワーク上で送信されるパケット内の記憶単位の総数をネットワーク上の所与のルートに関連する最大伝送単位（ＭＴＵ）と比較するステップと、
比較の結果がパケット内の記憶単位の総数がＭＴＵより大きいことを示している場合、さらに、
パケット内の記憶単位数をＭＴＵ記憶単位を保持できる１つまたは複数のパケット内に記憶された実質的に均等な単位のグループに分離することによってパケットを処理するステップとを実行することによって決定する命令シーケンスを表すコンピュータ・データ信号。

【請求項８】　所定のルート上で送信可能な最大伝送単位（ＭＴＵ）を受信するステップをさらに含む請求項７に記載の搬送波内に具体化されたコンピュータ・データ信号。

【発明の詳細な説明】

【０００１】

【発明の属する技術分野】本発明は、一般にネットワーク通信、より詳細に言えば、コンピュータの異種ネットワーク上でデータ・パケットを生成する方法およびシステムに関する。

【０００２】

【従来の技術】ネットワーク・コンピューティングは最近１０年間で著しい速度で発展した。ネットワーク・コンピューティング環境では、ユーザはネットワークに接続された複数のコンピュータにアクセスできる。ネットワーク・コンピューティングのトップ企業であるＳｕｎ　Ｍｉｃｒｏｓｙｓｔｅｍｓ　Ｉｎｃ．は、コンピューティング市場におけるこの発展する分野の商業的成功を強調すべく「Ｔｈｅ　Ｎｅｔｗｏｒｋ　Ｉｓ　ｔｈｅ　Ｃｏｍｐｕｔｅｒ」™というスローガンを中心に市場キャンペーンを張っているほどである（Ｔｈｅ　Ｎｅｔｗｏｒｋ　Ｉｓ　ｔｈｅ　Ｃｏｍｐｕｔｅｒ、Ｓｕｎ、Ｓｕｎのロゴ、Ｓｕｎ　Ｍｉｃｒｏｓｙｓｔｅｍｓ、Ｓｏｌａｒｉｓ、Ｕｌｔｒａ、Ｊａｖａは米国およびその他の諸国のＳｕｎ　Ｍｉｃｒｏｓｙｓｔｅｍｓ　Ｉｎｃ．の商標または登録商標。すべてのＳＰＡＲＣ商標は、米国およびその他諸国のＳＰＡＲＣ　Ｉｎｔｅｒｎａｔｉｏｎａｌ，Ｉｎｃ．のライセンスの下で使用され、その商標または登録商標。ＳＰＡＲＣ商標を付けた製品はＳｕｎ　Ｍｉｃｒｏｓｙｓｔｅｍｓ　Ｉｎｃ．が開発した

アーキテクチャに基づいている。ＵＮＩＸは米国およびその他の諸国の登録商標であり、Ｘ／Ｏｐｅｎ Ｃｏｍｐａｎｙ，Ｌｔｄ．を通じて排他的にライセンスされる。）。最近、数百万人のユーザがインターネットおよびワールド・ワイド・ウェブ上で利用できる数千のコンピュータにあるコンピューティング・リソースを開発するようになって、このスローガンは現実のものとなった。

【０００３】インターネットおよびワールド・ワイド・ウェブの普及は標準のネットワーク・プロトコルと異なるネットワークを結合するルータを使用したことによる。ＴＣＰ／ＩＰなどの通常のネットワーク・プロトコルはアプリケーション層、処理層、ホスト間プロトコル層（ＴＣＰ／ＵＤＰ）、インターネット・プロトコル層（ＩＰ）、ネットワーク・プロトコル層および物理層を含む。しばしばルータを使って物理層とネットワーク層にある異なるネットワークからの情報を変換している。特に、異なる物理層は、物理メディアを介して送信できる最大伝送単位（ＭＴＵ）を特に規定している該当するメディア・アクセス層（ＭＡＣ）を備える。一般に、ＭＴＵはすべてのネットワーク・セグメントが送信可能な最大送信単位の数として定義される。変換処理の一環として、ルータはＭＴＵが大きいネットワークから送信されるパケットをＭＴＵがより小さいネットワーク上のより小さいセグメントに断片化することを要求される。

【０００４】例えば、インターネットは異なる物理メディアと、イーサネット、ＩＥＥＥ８０２．３（ＣＳＭＡ／ＣＤ）、ＩＥＥＥ８０２．４（トークン・リング）、ＰＰＰ／ＳＬＩＰ（ダイヤルアップ）、Ｔ－１、Ｔ－３、およびＣＤＭＡ、ＴＤＭＡ、ＡＭＰＳ、ＧＳＭなどの無線技術の空中インタフェースを含むそれに対応するＭＡＣ層を結合する異種ＩＰネットワークである。上位層ではこれらの異なるネットワークはＴＣＰ、ＵＤＰなどのホスト間プロトコルおよびメール、フィンガ、およびｆｔｐなどのアプリケーション・プロトコルを共用する。異なるネットワーク間ですらこれらの上位層プロトコルは互換性がある。ただし、ネットワークおよびＭＡＣ層でこれらの異なるネットワークのＭＴＵ値（１９９０年１１月現在、インターネット上の様々なデータ・リンクに関連する一般的ＭＴＵ（バイト単位）には、以下のものがある。：６５５３５ Ｔｈｅｏｒｅｔｉｃａｌ Ｍａｘｉｍｕｍ（ＲＦＣ ７９１）；６５５３５ Ｈｙｐｅｒｃｈａｎｎｅｌ（ＲＦＣ １０４４）；１７９１４ ＩＢＭ Ｔｏｋｅｎ Ｒｉｎｇ（ＲＦＣ ７９１）；８１６６ ＩＥＥＥ ８０２．４（ＲＦＣ １０４２）；４４６４ ＩＥＥＥ ８０２．５（最高４Ｍｂ）（ＲＦＣ １０４２）；４３５２ ＦＤＤＩ（改訂版）（ＲＦＣ １１８８）；２０４８－４３５２ Ｗｉｄｅｂａｎｄ Ｎｅｔｗｏｒｋ（ＲＦＣ １１８８）；２００２ ＩＥＥＥ ８０２．５（４Ｍｂ推奨）（ＲＦＣ １０４２）；１５３６－２００２ Ｅｘｐ．ＥｔｈｅｒｎｅｔＮｅｔ（ＲＦＣ ８９５）；１５００ Ｅｔｈｅｒｎｅｔ Ｎｅｔｗｏｒｋ（ＲＦＣ ８９４）；１５００ Ｐｏｉｎｔ－ｔｏ－Ｐｏｉｎｔ（デフォルト）（ＲＦＣ １１３４）；１４９２ ＩＥＥＥ ８０２．３（ＲＦＣ １０４２）；１００６－１４９２ ＳＬＩＰ（ＲＦＣ １０５５）；１００６ ＡＲＰＡＮＥＴ（ＢＢＮ １８２２）；５７６－１００６ Ｘ．２５ Ｎｅｔｗｏｒｋ（ＲＦＣ８７７）；５４４ ＤＥＣ ＩＰ Ｐｏｒｔａｌ；５１２ Ｎｅｔｂｉｏｓ（ＲＦＣ １０８８）；５０８ ＩＥＥＥ ８０２／Ｓｏｕｒｃｅ－Ｒｔ Ｂｒｉｄｇｅ；（ＲＦＣ １０４２）；５０８ ＡＲＣＮＥＴ（ＲＦＣ １０５１）；２９６－５０８ Ｐｏｉｎｔ－ｔｏ－Ｐｏｉｎｔ（低遅延）（ＲＦＣ １１４４）；６８ Ｏｆｆｉｃｉａｌ ｍｉｎｉｍｕｍ ＭＴＵ．ＲＦＣ番号を使って、インターネット・エンジニアリング・タスク・フォース（ＩＥＴＦ）発行の対応する文書を見つけることができる。）は多種多様で、サイズの互換性がないことが多いデータグラム・セグメントを生成する。

【０００５】データグラムは接続レスのネットワーク上で送信される情報パケットであるため、ＭＴＵの非互換性の問題が生じる。換言すれば、その中で情報パケットの移動が制限される接続が送信元ノードと宛先ノードの間に存在しない。その代わりに、各パケットは通常異種ネットワークの異なるパス上で送信され、宛先ノードに到着する。その結果、各中継ネットワークのＭＴＵに合わせて各データグラムを異なる量で断片化することができる。例えば、イーサネット・ネットワークから送信されるデータグラム・セグメントは通常１５００バイト長であり、より小さいいくつかのデータグラム・セグメントに断片化してＸ．２５（ＭＴＵ５７６）またはＮｅｔｂｉｏｓ（ＭＴＵ５１２）物理ネットワーク層で送信される必要がある。このように、インターネットで用いられるルータはより小さなデータグラム・セグメント・サイズを必要とする、またはこれらのセグメントをネットワークの他の部分で互換性があるより小さいセグメントに断片化する大量のプロセッサ・サイクルを費やす。

【０００６】残念ながら、連続的な断片化によってネットワークの相互運用性を達成するとデータ処理および送信オーバヘッドが増加するためにネットワークのパフォーマンスは低下する。当初、大きいセグメントがますます小さいセグメントに縮小されてＭＴＵがますます小さくなる物理ネットワーク層を備えたネットワーク上で送信されるにつれて、断片化がルータの処理能力を消耗させる。受信側では、宛先ホストはより多くの処理能力とリソースを使って拡張されたセグメント・プールをバッファに入れてその拡張されたセグメント・プールを元のメッセージに組み直す。さらに、ネットワーク上で追加のパケットを送信すると送信されるパケットのヘッダ情

報の量が増え、有効なネットワーク帯域幅が減少する。

【０００７】断片化の制御は接続レス・ネットワーク上では特に困難である。それは、前述したように、各セグメントが異なるルートで送信できるからである。接続レス・ネットワーク上のデータグラムとして知られるいくつかのセグメントは大きいＭＴＵを備えたルート上を送信でき、断片化の必要はない。しかし、ネットワークのあるセグメントが障害になるか不正確に組立てられると、ネットワーク・トポロジが変更され、以降のセグメントはＭＴＵがより小さくまた広範囲の断片化を必要とするルート上で送信できる。複雑なルーティングによって送信元ホストは断片化を最小にする適切なデータグラム・サイズを選択することが困難になる。例えば、インターネット上で使用される一次ルートが８１６６バイトのＭＴＵを備えたＩＥＥＥ８０２．４ネットワークを含み、二次ルートが５７６バイトというはるかに小さいＭＴＵを備えたＸ．２５ネットワークを含むと仮定する。送信パケット数を最小化するため各データグラムが２つのルートの大きい方のＭＴＵに基づき、８１６６バイトを含むと仮定する。当初、一次ルート上で送信されるデータグラムは断片化を必要としない。しかし、一次ルートが使用できなくなるか話中の場合、ネットワーク・トポロジが変更されて５７６バイトというより小さいＭＴＵを備えた二次ルートがトラフィックを引き受ける。したがって、ルータは残りのデータグラムのそれぞれを断片化し、上述したように有効なネットワーク帯域幅を縮小するインターネット上で送信する。接続レス・ネットワーク上で断片化を最小にするデータグラム・サイズを選択するのはまさしく困難な作業である。

【０００８】断片化全体を減らす一方でネットワーク帯域幅を最適化するデータグラム・サイズを選択する現在の技法は、まだ十分に成功していない。一般にはデータはローカル・ネットワーク上で送信されローカル・ネットワークＭＴＵに基づいてデータグラムを生成する。残念ながらこの技法では、ローカル・ネットワークのＭＴＵサイズより小さいＭＴＵを備えたリモート・ネットワーク上でデータグラムが送信される場合は断片化が実行され、オーバヘッドが増加する。例えば、データがルータ装置とＸ．２５ネットワークで結合されたあるイーサネットから別のイーサネットへ送信されると仮定する。さらにイーサネット・ネットワークのＭＴＵが１５００バイト、２つのイーサネット・ネットワーク間のＸ．２５ネットワークのＭＴＵが５７６バイトと仮定する。すると、この第１の技法を使って生成されたデータグラムは、Ｘ．２５ネットワークをまたがる２つの５７６バイトのデータグラムと１つの３４８バイトのデータグラムに断片化される。この技法は通常データグラムがローカル・ネットワークにある限り断片化を減少させない。

【０００９】別の技法は、送信元ホストとして同じローカル・ネットワークにもサブネットにも接続されていないネットワークに送信する場合に、小さい方の５７６バイトすなわちファーストホップＭＴＵに基づいてデータグラムの大きさを決定する。この場合、あるパケットが通過したルータ、スイッチ、またはブリッジの数を数えることでホップ・カウントがネットワーク上の移動距離を表す。あるパケットに関連付けられたホップ・カウントＭはパケットがおよそＭ個の異なるネットワークを通過したことを意味する。当然の結果として、ネットワークのファーストホップＭＴＵは送信側ネットワークからちょうど１ホップ分離れている。残念なことに、ＭＴＵを選択するこの技法は通常必要な数に足りないデータグラムを生成しネットワーク機能が低下して使用可能なネットワーク帯域幅を浪費する。さらに、データグラムが５７６未満のＭＴＵを備えたネットワークを通過すると断片化が行われる。

【００１０】第３の技法はネットワーク上の送信元ノードと宛先ノードの間のルートのパスＭＴＵ（ＰＭＴＵ）に基づいてデータグラムのサイズを決定する。ＩＰネットワークに適用されるＰＭＴＵ検出技法は、Ｊｅｆｆｒｅｙ ＭｏｇｕｌとＳｔｅｖｅ Ｄｅｅｒｉｎｇ著の、インターネット特別技術調査委員会（ＩＥＴＦ）に関連して出版された「Ｐａｔｈ ＭＴＵ Ｄｉｓｃｏｖｅｒｙ」という題のＲＦＣ１１９１に記載されている。要するに、この技法はＩＰヘッダの「Ｄｏｎ’ｔＦｒａｇｍｅｎｔ」（ＤＦ）ビットを使って任意のネットワーク・パスのＰＭＴＵを動的に発見する。最初、送信元ホストはＰＭＴＵをファーストホップＭＴＵと仮定してＤＦビットを立ててそのパス上で全データグラムを送信する。これらのデータグラムが任意のネットワーク・パスで何らかのルータによって断片化なしに転送するには大きすぎる場合、当該ルータはサイズ超過のデータグラムをドロップし送信元ノードに「断片化が必要でＤＦビットはオン」を示す特別のコードを付けたインターネット制御メッセージ・プロトコル（ＩＣＭＰ）デスティネーション・アンリーチャブル（Ｄｅｓｔｉｎａｔｉｏｎ Ｕｎｒｅａｃｈａｂｌｅ）メッセージを返送する。これに応答して、送信元ノードはデータグラムが断片化なしに送信できるまでその大きさをますます小さくして上記の処理を繰り返す。残念なことに、インターネットなどの接続レス・ネットワーク上で送信されるそれぞれの連続的なデータグラムは別のルートを通るため、同じネットワーク上で送信された以前のデータグラムとは異なるＰＭＴＵ値を備える。このように、Ｐａｔｈ ＭＴＵ Ｄｉｓｃｏｖｅｒｙを用いたシステムでさえ断片化を減らすことができず、したがってパケット送信量が増えるとネットワークの有効帯域幅が減ることになる。

【００１１】

【発明が解決しようとする課題】したがって、必要なものは断片化の量を削減し、ネットワーク上でのパケット送信に関連付けられた有効ネットワーク帯域幅を拡張す

るネットワークで用いるパケット生成の方法およびシステムである。

【００１２】

【課題を解決するための手段】本発明はネットワーク上の異なるルートで送信するパケットを生成する改良された方法およびシステムを提供する。インターネットなどの巨大なネットワークでは、ネットワーク上の各ルートはより小さい相互運用が可能なセグメントに断片化されるまではネットワーク間で直ちには送信できない異なるサイズのパケットを送信する。まず、この技法は所定のルートで送信可能な最大送信単位（ＭＴＵ）を決定する。次に、ネットワーク上で送信する各パケットのサイズがＭＴＵのサイズと比較される。この比較でパケットが所定のＭＴＵより大きいとわかると、パケットはさらに処理されてからルート上で送信される。さらに別の処理ではまずパケットに含まれる全送信単位数をＭＴＵ値で割る。この除算の結果はＤＣＯＵＮＴ記憶ユニットに一時記憶され、除算の余りはＲＣＯＵＮＴ記憶ユニットに記憶される。ＲＣＯＵＮＴの余りが非ゼロの場合、ＤＣＯＵＮＴの値が１つ増分される。ＤＣＯＵＮＴ値は現在の技法を用いてパケットを送信する場合の最小データグラム数を示す。次に、元のパケットに含まれる送信単位はＤＣＯＵＮＴパケットに公平に分配され、ネットワーク・ルートでの送信用に準備される。

【００１３】本発明は従来技術では使用できなかったいくつかの利点がある。まず、本発明で提供される技法は、送信されるデータグラム数が実質的に変更されないためネットワーク・オーバヘッド量を増やさずに断片化の確率を減らす。最良のケースではこれらの技法は大幅に断片化を減らして有効ネットワーク帯域幅を拡張するが、最悪のケースではこれらの技法の有効性は他の技法と同じ程度である。

【００１４】本発明はネットワーク・プロトコルに非互換性を持ち込まずにネットワーク・パフォーマンスを向上させるため、また有利である。大半の大きいネットワークではネットワーク・プロトコル内の低レベル層の変更は不可避的に非互換となりネットワークの相互運用性が減少する。これはＩＰやＴＣＰ／ＩＰなどのプロトコルを実行する数百、数千のコンピュータが直ちに変更できないインターネットについて当てはまる。本明細書に記載したネットワーク処理の改良は基本的にＩＰプロトコル動作を変更しないため、本発明を使用してもこのジレンマは発生しない。したがって、本発明の教示によってネットワークの相互運用性を維持し、データ・パケットを送信するために有効なネットワーク帯域幅を改善するという困難な問題は明快に解決される。

【００１５】

【発明の実施の形態】図１に本発明の一実施形態を実施するコンピュータ・ネットワーク１００を示す。この例では、送信元ノード１０２と宛先ノード１１４がルータ１０６および１１０を使ってネットワーク１０４、１０８、および１１２を含む一次ルート上で相互にデータを送信する。また別に、送信元ノード１０２および宛先ノード１１４はルータ１０６および中間ノード１１６を使ってネットワーク１０４、１１８、および１１２を含む二次ルート上で通信を行うことができる。コンピュータ・ネットワーク（図１）は通常他のデータ装置へ１つまたは複数のネットワークを介してデータ通信を行う。例えば、ネットワーク１００は送信元ノード１０２および宛先ノード１１４の間に一部現在「インターネット」という通称があるワールドワイドなパケット・データ通信ネットワークを使ってネットワーク接続を行うことができる。インターネットは電気、電磁気、および光信号を使ってさまざまなタイプの情報を表すディジタル・データ・ストリームを伝送する。コンピュータ・ネットワーク１００を介して伝送される信号は送信元ノード１０２への、また送信元ノードからのディジタル・データを伝送し、情報を送信する搬送波の代表的な形式である。

【００１６】送信元ノード１０２はネットワーク１００を介してメッセージを送信し、プログラム・コードを含むデータを受信する。インターネットの例では、宛先ノード１０４はアプリケーション・プログラムの要求コードをネットワーク１００を使ってインターネット経由で送信元ノード１０２に送信できる。本発明は、その種のダウンロードされたアプリケーションの１つが異種ネットワーク上でデータ・パケットを生成する方法とシステムであり、本明細書に記載される。受信されたコードは受信時に送信元ノード１０２によって実行され、または記憶装置に記憶されて後ほど実行される、あるいはその両方の処理がされる。このように、送信元ノード１０２は搬送波の形式でアプリケーション・コードを入手できる。

【００１７】上記のルータおよびノードは汎用コンピュータの内部で実行される処理として、またはネットワーク上で送信されるデータ・パケットを受信して処理するための専用のスタンドアロン装置として実施できることは当業者は理解するであろう。これに応じて、これらの装置の一般のコンポーネントを図２に示す。

【００１８】図２について説明する。同図には、ネットワーク１００（図１）上のルータまたはノードとして動作するよう構成されたコンピュータ・システムの必須コンポーネントが示されている。図２のコンピュータ・システム２０２は第１のネットワーク・インタフェース２１２、第２のネットワーク・インタフェース２１１、プロセッサ２１４、一次記憶２１６、二次記憶２１８、および前記の要素間の通信を容易にする入出力インタフェース２２０を含む。ネットワーク・インタフェース２１２はコンピュータ・システム２０２をネットワーク１００（図１）に結合し、コンピュータ・システム２０２とネットワーク１００（図１）上のその他のノード間の通

信を容易にする。ネットワーク・インタフェース２１１はコンピュータ・システム２０２を第２のネットワーク（図示せず）に結合し、コンピュータ・システム２０２とネットワーク１００（図１）上のその他のノードまたは第２のネットワーク（図示せず）上のノード間の通信を容易にする。ルータ装置は通常、複数の異なるネットワークを結合するために、複数のネットワーク・インタフェースを備えていることは当業者には周知である。したがって、図２の代表的なコンピュータ・システム２０２は２つの異なるネットワーク用に構成されているが、１つまたは複数のネットワーク・インタフェースを追加することで別のネットワークに結合できる。

【００１９】図２で、コンピュータ・システム２０２内のプロセッサ２１４は入出力インタフェース２２０を介して一次記憶２１６からのコンピュータ命令を取り出す。これらの命令を受信すると、プロセッサ２１４はこれらのコンピュータ命令を実行する。これらのコンピュータ命令を実行すると、プロセッサ２１４は一次記憶２１６のデータを検索し、記憶に書き込み、１つまたは複数のコンピュータ・ディスプレイ装置（図示せず）に情報を表示し、１つまたは複数の入力装置（図示せず）からコマンド信号を受信し、二次記憶２１８または送信元ノード１０２、宛先ノード１１４、中間ノード１１６、ルータ１０６、またはルータ１１０などのネットワーク１００（図１）上のその他のノードのデータを受信しあるいはデータを書き込むことができる。プロセッサ２１４はまた、図２には図示していない。コンピュータ・システム２０２に結合された他のネットワークに関するノードに対して上記の機能を実施することができる。一次記憶２１６および二次記憶２１８は、ランダム・アクセス・メモリ（ＲＡＭ）、読出し専用メモリ（ＲＯＭ）、磁気記憶装置およびＣＤ－ＲＯＭなどの光記憶メディアを含めてあらゆる種類のコンピュータ記憶を無制限に含むことができることを当業者はまた理解する。プロセッサ２１４はカリフォルニア州Ｍｏｕｎｔａｉｎ ＶｉｅｗのＳｕｎ Ｍｉｃｒｏｓｙｓｔｅｍｓ，Ｉｎｃ．製のＳＰＡＲＣ互換プロセッサ、ＵｌｔｒａＳｐａｒｋ互換プロセッサまたはＪａｖａ互換プロセッサのいずれかである。別に、プロセッサ２１４はカリフォルニア州ＣｕｐｅｒｔｉｎｏのＡｐｐｌｅ，Ｉｎｃ．製のＰｏｗｅｒＰＣプロセッサ、Ｉｎｔｅｌ ＣｏｒｐｏｒａｔｉｏｎまたはＡＭＤ、およびＣｙｒｉｘ製のすべてのＰｅｎｔｉｕｍまたはｘ８６互換プロセッサ、または他のすべての特殊用途のプロセッサに基づくことができる。

【００２０】図２について説明する。同図では、一次記憶２１６がコンピュータ・リソースを管理するオペレーティング・システム２２２を含む。このオペレーティング・システムは、Ｓｏｌａｒｉｓオペレーティング・システムか、ネットワーク・コンピューティング環境でのデータ・パケットおよびデータグラムの処理に関連付けられた処理要件をサポートする能力があるすべてのオペレーティング・システムであることが好ましい。別の実施形態では、オペレーティング・システム２２２はＣＩＳＣＯ Ｉｎｃ．のルータ装置に使われるインターワーキング・オペレーティング・システム（ＩＯＳ）またはＡｓｃｅｎｄ Ｉｎｃ．、Ｂａｙ Ｎｅｔｗｏｒｋ，Ｉｎｃ．、または３Ｃｏｍ，Ｉｎｃ．のルータ装置に使われる類似のオペレーティング・システムである。また一次記憶にはＴＣＰ／ＩＰ、Ｘ．２５、ＳＮＡなどの１つまたは複数のプロトコル・スタック２２４、または異なるネットワーク間でデータグラムを変換するＮｅｔＷａｒｅ（ＮｅｔＷａｒｅは米国およびその他諸国のＮｏｖｅｌｌ，Ｉｎｃ．の登録商標）などのネットワーク・オペレーティング・システムの一部も含まれる。本発明の一実施形態ではこれらのプロトコル・スタック２２４と、データグラム生成に用いるデータグラム生成器２２６が連動している。

【００２１】実際、図２に示すルータ２０２は異なるネットワーク間でデータグラムを変換するのに必要なプロトコル・スタックの一部しか含まない場合がある。例えば、ＩＰネットワーク内ではルータは通常２つの異なるネットワークでは異なるＩＰ層の下の層を全て含む。図３について説明する。同図で、ルータ装置３０６は、第１のネットワーク上で使用される下位レイヤ・メディア・アクセス・プロトコル１（ＭＡＣ１）と第２のネットワーク上で使用されるメディア・アクセス・プロトコル２（ＭＡＣ２）に加えてＩＰ層を含む。上述したように、異なるネットワーク・メディアはそれに対応する異なるＭＡＣプロトコルを利用する。一方、異なるネットワーク上の各ノードはネットワーク・アクセス層、トランスポート層、およびアプリケーション層を含むアプリケーションを処理するプロトコル・スタック一式を含む。これらの追加の層はエンド・ユーザ、アプリケーション、またはリモート装置に上位レベル機能を提供する。これに合わせて、図３では、ノード３０２および３１０がＩＰ層および該当するメディア・アクセス・プロトコルに加えてアプリケーション層とＴＣＰ層を含むことが示される。

【００２２】この時点で、異種コンピュータ・ネットワークでのＩＰ層の役割の基本説明が本発明の舞台を設定する上で役立つであろう。図４ＡはＩＰ層を使ってＸ．２５パケット交換ネットワーク４０７上の第１のノード４０２からイーサネット・ローカル・エリア・ネットワーク４１２上の第２のノード４１４へデータを送信する方法を示すブロック図である。各プロトコル・スタックの管理と、送信されたデータの処理に関連付けられたハードウェアは、この例ではＩＰプロトコルの動作を強調するために省略している。図４Ｂは図４Ａに示すＩＰネットワーク上で送信するデータ・パケットへの変更をステップ１～１５に示す図である。図４Ｂの各セグメント

は簡潔に次の略号で示す。トランスポート・ヘッダ（Ｔ－Ｈ）、インターネット・プロトコル・ヘッダ（ＩＰ－Ｈ）、Ｘ．２５パケット・ヘッダ（Ｐ－Ｈ）、ＬＡＰ－Ｂリンク・ヘッダ（Ｌ－Ｈ）、ＬＡＢ－Ｂトレーラ（Ｌ－Ｔ）、ＬＬＣヘッダ（ＬＬＣ－Ｈ）、ＭＡＣヘッダ（ＭＡＣ－Ｈ）、ＭＡＣトレーラ（ＭＡＣ－Ｔ）。

【００２３】図４Ａについて説明する。同図では、第２のノード４１４がノード４０２に返送するデータがＴＣＰ層に対応するＩＰヘッダおよびトランスポート・ヘッダを備えたデータグラムであるインターネット・プロトコル・データ単位にカプセル化される。ＩＰネットワーク上でデータグラムを送信するため、データグラムＩＰヘッダは第１のノード４０２のアドレスを含み、下位レベル・ネットワーク・パケット・ヘッダはルータ４１０のＭＡＣアドレスを含む。ルータ４１０のＭＡＣアドレスは第１のノード４０２のアドレスと共に含まれるが、これは第１のノード４０２から直接第２のノード４１４にアクセスできないためである。ルータ４１０は、イーサネット・ネットワーク４１２およびＸ．２５ネットワーク４０７で使われる、異なるネットワーク・プロトコルの間の連絡機能として動作する。この例では、Ｘ．２５パケット交換ネットワーク層３を使ってＩＰパケットがＸ．２５ネットワーク上で移動する際にカプセル化している。最終的に、カプセル化されたデータグラムはＸ．２５ネットワーク４０７を介して第１のノード４０２に送信され、ここで元のＩＰデータグラムが使用できるようになり処理が可能となる。ここではＸ．２５ネットワーク動作の詳細は本明細書の範囲外であるが、Ｘ．２５ネットワークによって元のＩＰデータグラムが変更されないことに注意するべきである。

【００２４】実際にＩＰデータグラムを送信する前に、ルータ４１０はＩＰデータグラムがデータグラムを受信するネットワークのＭＴＵを超えないことを確認する必要がある。例えば、データグラム・セグメントが受信側ネットワークのＭＴＵより大きい場合、ルータ４１０はそのデータグラムを１つまたは複数のより小さいデータグラムに断片化する必要がある。図４Ａで、Ｘ．２５上で送信されるＭＴＵが５７６のデータグラムはＭＴＵが1500バイトのイーサネット・ネットワーク（ＩＥＥＥ８０２．３）から受信した場合、少なくとも３つのデータグラムに分割する必要がある。上述したように、データグラムが小さくなると送信されるヘッダ情報の割合が増え、通常はパケットの再送数が増えるため、断片化によってネットワークの効率は低下する。

【００２５】過去において、断片化を抑えるための技法は適切なデータグラムのサイズに焦点を当てていたが、実際のデータグラム内にデータがどのように分配されるかを見過ごしていた。最も一般的な技法はデータグラムを順次データで満たし、結果的に一連の大小のデータグラムが出来上がる。図５Ａおよび図５Ｂに、従来技術でＭＴＵが１５００バイトのネットワークで４７００バイトのデータグラムと１７００バイトのデータグラムを複数のデータグラムに不均等に断片化する処理を示す。この方法では、図５Ａの４７００バイトのデータグラムはより小さい１５００バイトのデータグラム３つと４番目の２００バイトのデータグラムに分配され、図５Ｂの１７００バイトのデータグラムはより小さい１５００バイトのデータグラム２つと２００バイトのデータグラム１つに分配される。データグラムが比較的大きいとＭＴＵがそれより小さいネットワークに遭遇することが多いため、従来技術では頻繁な断片化は珍しくなかった。

【００２６】これとは対照的に、本発明の実施形態はこれらのデータグラム間に均等にデータを分配し、最大断片化サイズを減らして断片化の確率を減らす技法を提供することである。ＩＰデータグラムを生成するこの新しい技法によるデータの均等な分配の例が図５Ｃおよび図５Ｄである。この方法では、本発明の原理によって分割された４７００バイトのデータグラムはそれぞれが１１７０バイトのデータを備えた４つのデータグラムに分配され、図５Ｄの１７００バイトのデータグラムはより小さい８５０バイトのデータグラム２つに分配される。この明快な解決法によって、ネットワークの非互換性を持ち込まずにデータグラム生成に変更を加えることでネットワーク効率は向上する。その結果、断片化されずにネットワークを通過できるデータグラムの数は増える。以下に、ＩＰネットワークに関して本発明の一実施形態によるデータグラム生成を詳述する。

【００２７】まず、本発明はネットワークのＭＴＵに決定技法があることを前提とする。通常、ＭＴＵは静的または動的決定方法を使って決定される。いずれの方法でも、本発明は断片化を減らすことでネットワーク帯域幅を向上させることができる。一実施形態では、ネットワークのＭＴＵは静的に決定される。この場合、データを送信する送信元ホストはＭＴＵの値を合理的な程度に小さいＭＴＵ値、例えば５７６バイトに統計的に設定するか、これとは別に送信元ホストはＭＴＵを送信元ホストとして同じローカル・ネットワークまたはサブネットに接続していないすべてのネットワークへのファーストホップＭＴＵと５７６バイトのうち小さい方の値に設定する。

【００２８】別の実施形態では、送信元ホストは所定の時間間隔でのインターネットの所与のルートに沿ってノードに問い合わせを行って、あるパスのＭＴＵを動的に決定してパスＭＴＵ（ＰＭＴＵ）を決定する。ＰＭＴＵは所与の特定のルートでの各ホップに関連付けられたＭＴＵの最小値を示し、ＩＰネットワークで送信されるデータグラムのＭＴＵの上限値となる。インターネットでは、ＰＭＴＵは「Ｐａｔｈ　ＭＴＵ　Ｄｉｓｃｏｖｅｒｙ」と呼ばれる技法を使って利用でき、Ｊｅｆｆｒｅｙ　ＭｏｇｕｌとＳｔｅｖｅ　Ｄｅｅｒｉｎｇ著の、イン

ターネット特別技術調査委員会（ＩＥＴＦ）に関連して出版された「Ｐａｔｈ　ＭＴＵ　Ｄｉｓｃｏｖｅｒｙ」という題のＲＦＣ１１９１に記載されている。ＰＭＴＵを実施する場合、本発明は、ネットワークでデータグラムが送信される際にホスト間でネットワーク状態およびローカル・ホスト・リソースに関する情報を交換できるホスト間通信機能が存在することをさらに前提としている。インターネットおよびその他のＩＰベースのネットワークで通常使われるホスト間データグラム通信の１つのタイプにインターネット制御メッセージ・プロトコル（ＩＣＭＰ）があり、Ｊ．Ｐｏｓｔｅｌ著のＩＥＴＦに関連して出版された「Ｉｎｔｅｒｎｅｔ　Ｃｏｎｔｒｏｌ　Ｍｅｓｓａｇｅ　Ｐｒｏｔｏｃｏｌ」という題のＲＦＣ７９２に記載されている。Ｐａｔｈ　ＭＴＵ　ＤｉｓｃｏｖｅｒｙおよびｉＣＭＰに関する上記の両文書を本発明の一実施形態で用いている。ＩＣＭＰおよびＰＭＴＵＤｉｓｃｏｖｅｒｙに類似したネットワーク・サービスを有するネットワーク上での作業に本発明の代替実施形態を適用できることを当業者は理解するであろう。

【００２９】図６について説明する。同図は本発明の一実施形態によるデータグラム・パケットの最適サイズを決定するためのステップのすべてを含む流れ図である。まず、ステップ６０２で、本方法はインターネット上の送信元ホストと宛先ホストの間の所与のルートの最大伝送単位（ＭＴＵ）を決定する。前述したように、ＭＴＵの選択は時に困難な処理である。大半の接続レス・ネットワークでは、ＭＴＵ値はルーティング・トポロジの変化に伴って急速に変動し、データグラムは異なるルートで宛先ホストに送信される。一次ルートが話中または物理的に切断されて使用不可になると、ネットワーク・トポロジが変化してパケットは別の二次ルートへルーティングされなくてはならない。通常、ルートはそれぞれ別々のＭＴＵ値を備えている。一実施形態では、上記のＰＭＴＵ検出技法はデータ送信を行うたびにそれに先立って一度実行され、ＭＴＵが決定される。本発明により、このＭＴＵ値をデータグラム生成開始時点に使うことができる。また、本発明の別の実施形態では、データ送信中に所定の時間間隔で一連のＰＭＴＵ検出を実行して、時が経つと発生する可能性があるＭＴＵサイズの変動に対処することができる。一般に、これらのＰＭＴＵ検出の周期はネットワークのパフォーマンス全体に影響が出るほど短くしてはならない。上記のＲＦＣ１１９１　Ｐａｔｈ　ＭＴＵ　ＤｉｓｃｏｖｅｒｙおよびＲＦＣ７９２　ＩＣＭＰは、ＩＣＭＰおよびＰａｔｈ　ＭＴＵ　Ｄｉｓｃｏｖｅｒｙを併用して所与のルートのＭＴＵを決定する方法を簡潔に述べている。本発明の実施形態でも上記の静的ＭＴＵ決定技法を使って所与のネットワークのＭＴＵ値を決定できるのはもちろんである。

【００３０】次に、判定ステップ６０４で本方法は現在のデータグラムのサイズを上記のように決定したＭＴＵ値と比較する。現在のデータグラムのサイズがＭＴＵ以下であれば、データグラムはそれより小さいデータグラムに断片化しなくても送信元ホストから宛先ホストへ送信できる程度に小さい。その場合、ステップ６０８でさらにデータグラム処理を行うためにデータグラムが送信される。しかし、データグラムのサイズがＭＴＵ値より大きい場合は、以下のステップでデータグラムをその後の断片化の確率が最小化される一連のより小さいデータグラムに分割する。

【００３１】図６のステップ６１０で、ＭＴＵ値と元のデータグラムまたはデータ・パケットの総バイト数を使っていくつのより小さいデータグラムを送信する必要があるかが決定される。ステップ６１０で、元のデータグラムに含まれる総バイト数が上で決定したＭＴＵ値で割られる。この除算の整数結果は一時可変ＤＣＯＵＮＴに記憶され、余りがあれば一時可変ＲＣＯＵＮＴに記憶される。これらの値は判定ステップ６１２で使われて元のデータグラムのデータを伝送するために追加のデータグラムが必要かどうか判定される。判定ステップ６１２で、ＲＣＯＵＮＴに記憶された余りがゼロ値と比較される。比較の結果、ＲＣＯＵＮＴが非ゼロの場合、処理はステップ６１４へ移行し、ＤＣＯＵＮＴ値を増分して追加のデータグラムが必要であることが示される。ただし、ＲＣＯＵＮＴ値がゼロの場合、元のデータグラム内のデータを送信するために追加のデータグラムは必要でなく、処理はＤＣＯＵＮＴを増分せずにステップ６１６へ移行する。

【００３２】ステップ６１６で、元のデータグラム内のバイト数がＤＣＯＵＮＴに格納されている数だけのいくつかのデータグラムに均等に分配される。生成された小さいデータグラムはそれぞれ特定の断片に合わせてアドレス、オフセット、チェックサム、およびその他の特定の値が調整された状態で元のデータグラムに含まれる同じヘッダ情報を受信することを当業者は理解するであろう。従来技術によって生成されたデータグラムとは対照的に、これらのデータグラムはデータグラム値に大小のばらつきがない。その代わりに、バイトはより小さいデータグラム・セグメントのそれぞれに均等に分配される。こうしてできたデータグラムはサイズがほぼ等しく、ＭＴＵが小さいルートでデータグラムが送信される際に断片化される可能性はより低い。

【００３３】例として、図１に送信元ノード１０２と宛先ノード１１４の間に２つのルートがあるネットワークを示す。ルートＡはネットワーク１０４、１０８、および１１２を含み１１００バイトのパスＭＴＵ（ＰＭＴＵ）を備え、ルートＢはネットワーク１０４、.１１８、および１１２を含み１００６バイトのパスＭＴＵ（ＰＭＴＵ）を備える。図５Ｂにその種のネットワーク上で送信するために従来技術で１７００バイトのデータグラムを２つのより小さいデータグラムに分割する方法を示

す。図５Ｂの第１のデータグラムはルートＡ（ＰＭＴＵ１１００）を通る場合にルータ１０６を通ることは可能だが、さらに断片化をしないとルータ１１０経由では不可能である。同様にルートＢ（ＰＭＴＵ１００６）では、第１のデータグラムはルータ１０６からノード１１６を通るのに断片化が必要である。図５Ｂの第２のデータグラムはさらに断片化をしなくてもいずれのルートも通ることができる。さらなる断片化に関連付けられた追加のオーバヘッドの結果、オーバヘッドが発生してネットワーク効率は低下する。

【００３４】対照的に、図５Ｄは本発明の各実施形態が同じ数のバイトを同じ数のデータグラムにより最適な方法で分割する様子を示す。いくつかのデータグラム内のバイト数を増やして他のデータグラムのバイト数を減らすのではなく、本発明は送信データを最小数のデータグラムに均一に分割する。この例では、図５Ｄの両方のデータグラムはわずか８５０バイトしか含んでいない。したがって、これらのデータグラムは両方とも断片化なしに図１のルートＡとルートＢを通過できる。

【００３５】本発明の実施形態には当業界で以前は不可能だったいくつかの有利な点がある。第１に、本発明で提供された技法では送信されるデータグラム数は実質的に変わらないため、ネットワーク・オーバヘッドの量を増やさずに断片化の確率が低下する。最良のケースではこれらの技法は大幅に断片化を減らして有効ネットワーク帯域幅を拡張するが、最悪のケースでもこれらの技法の有効性は他の技法と同じ程度である。

【００３６】本発明の実施形態はネットワーク・プロトコルに非互換性を持ち込まずにネットワーク・パフォーマンスを向上させるため、また有利である。大半の大きいネットワークではネットワーク・プロトコル内の低レベル層の変更は不可避的に非互換となりネットワークの相互運用性が減少する。これは特にＩＰやＴＣＰ／ＩＰなどのプロトコルを実行する数百、数千のコンピュータが直ちに変更できないインターネットについて当てはまる。本明細書に記載したＩＰ層の改良は基本的にＩＰプロトコル動作を変更しないため、本発明を使用してもこのジレンマは発生しない。したがって、本発明の教示によってネットワークの相互運用性を維持し、有効ネットワーク帯域幅を改善するという困難な問題は明快に解決される。

【００３７】以上、本明細書で特定の実施形態について例示してきたが、本発明の精神と範囲を逸脱することなしにさまざまな修正を加えることができる。本発明の実施形態は多様な異なるネットワーク・プロトコルを使って実施でき、ＴＣＰ／ＩＰプロトコルを使用するネットワークに接続したコンピュータ・システムに限定されないことを当業者は理解するであろう。好ましい実施形態に実質的に類似した代替実施形態をＳＮＡ（システム・ネットワーク・アーキテクチャ）、ＩＰＸ／ＳＰＸ、Ａｐｐｌｅｔａｌｋ、またはＸ．２５などのその他のネットワーク・プロトコルを使って実施できる。ＴＣＰ／ＩＰおよびＳＮＡネットワークの統合に関する詳細については、Ｔａｙｌｏｒ著「Ｉｎｔｅｇｒａｔｉｎｇ ＴＣＰ／ＩＰｉｎｔｏ ＳＮＡ」、Ｗｏｒｄｗａｒｅ Ｐｕｂｌｉｓｈｅｒｓ、１９９３年を参照されたい。さらに、好ましい実施形態に実質的に類似した別の代替実施形態をＰＭＴＵ値が一回ただ収集されるだけではなく所定の時間間隔で無数の回数更新されるという点を除いて実施できる。好ましい実施形態に実質的に類似したさらに別の代替実施形態を最大パケット・サイズ送信制限があるすべてのネットワークに適用できるという点を除いて実施できる。本発明の例では一般にバイト数を使ってデータグラムのサイズを測定しているが、データグラムのサイズはいかなる他の伝送単位の尺度、例えば、時間間隔が測定するビット数、ブロック数、またはデータ・ストリーム数によってさえ測定できることを当業者は理解するであろう。

【００３８】これらの実施形態は、コンピュータを使用するものであるので、当然ソフトウエアを含み、それらはＣＤＲＯＭその他のコンピュータが読むことのできる記録媒体にコードとして具体化して記録される。その媒体に記録されるコードは、それぞれのルートが異なるサイズのパケットを伝送できるネットワーク上の異なるルートで送信されるパケットを生成できるように記録されている。さらに、そのコードは、ネットワーク上で送信されるパケット内の記憶単位の総数をネットワーク上の所与のルートに関連する最大伝送単位（ＭＴＵ）と比較するように構成された第１のコード部分と、比較の結果がパケット内の記憶単位の総数がＭＴＵより大きいことを示している場合にパケットをさらに処理する第２のコード部分と、パケット内の記憶単位数をＭＴＵ記憶単位を保持できる１つまたは複数のパケット内に記憶された実質的に均等な単位のグループに分離するように構成された第３のコード部分とを備えている。

【００３９】以上のように、本発明は上述した実施形態に限定されず、むしろ同じ価値のある物の完全な範囲にてらして添付の特許請求の範囲によって定義される。

【図面の簡単な説明】

【図１】　本発明の一実施形態を実施するコンピュータ・ネットワークを示す図である。

【図２】　ネットワーク上のルータまたはノードとして動作するように構成されたコンピュータ・システムの必須コンポーネントを示すブロック図である。

【図３】　コンピュータ・ネットワーク上の例示的なルータ装置およびノードで用いられる異なるプロトコル層を示すブロック図である。

【図４】　ＩＰ層を使ってパケット交換ネットワークの送信元ノードから宛先ノードへデータを送信する方法を示すブロック図（Ａ）とＩＰネットワークで送信される

データ・パケットの代表的な変更を示す図（B）である。

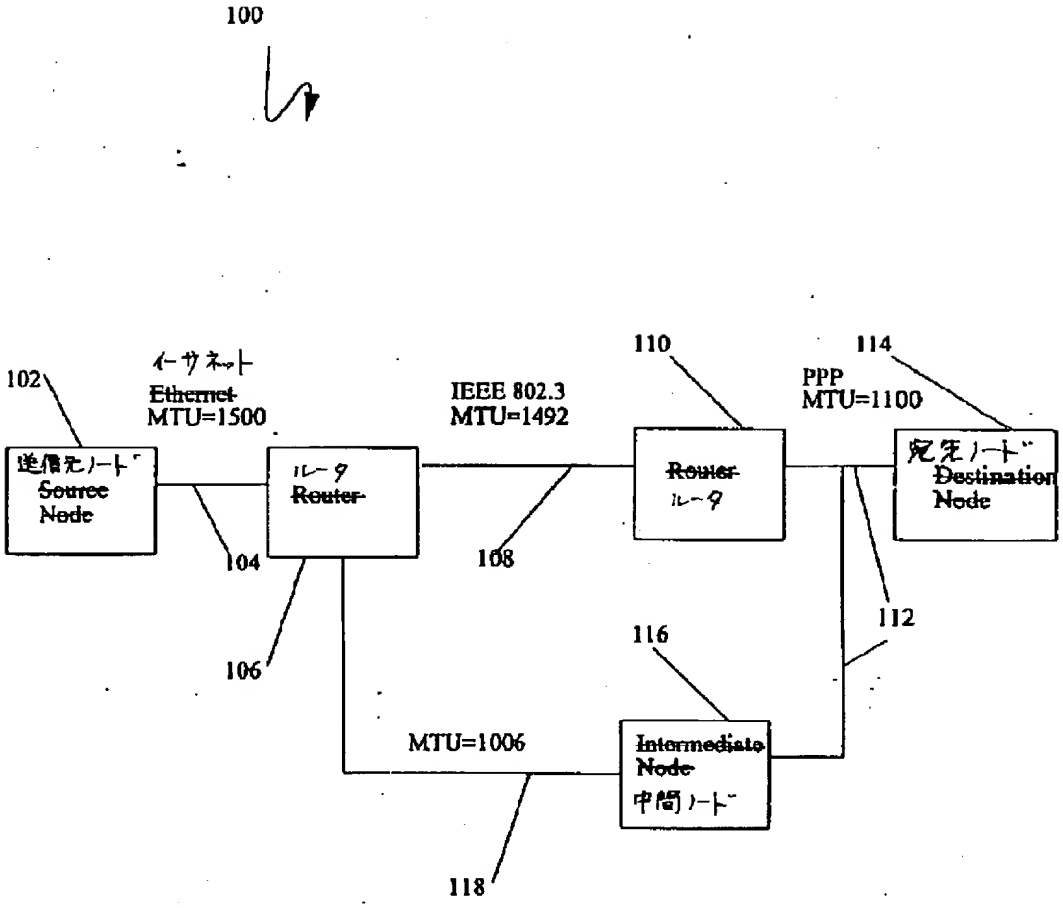
【図５】　代表的な従来技術の方法が大きいデータグラムをいくつかのデータグラムに不均等に断片化する方法を示す図（A，B）と本発明の一実施形態を使って大きいデータグラムをいくつかのデータグラムに均等に断片化する方法を示す図（C、D）である。

【図６】　本発明の一実施形態にしたがってデータグラム・パケットを生成するために用いられる全ステップを示す流れ図である。

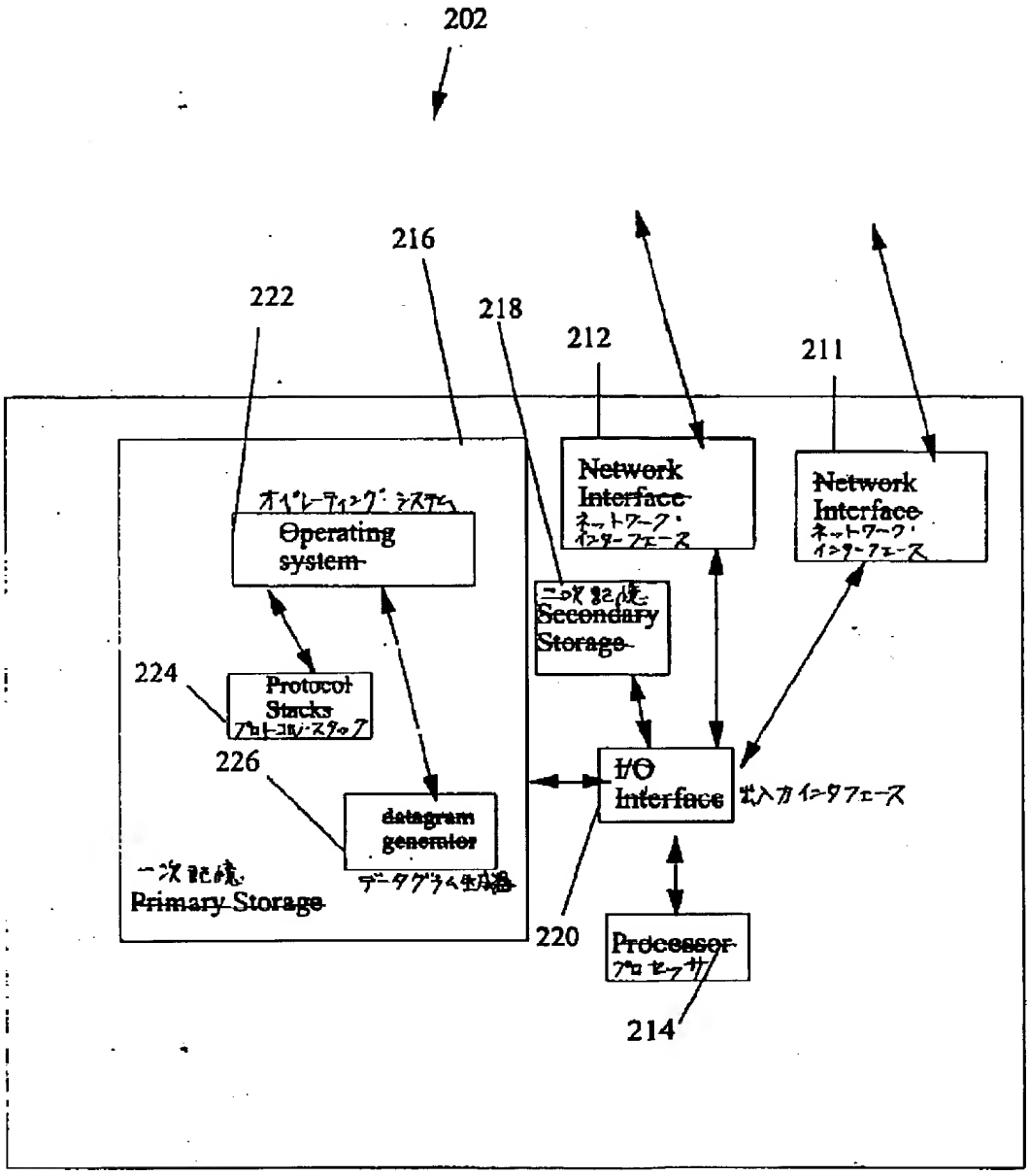
【符号の説明】

２０２　コンピュータ・システム
２１１　第２のネットワーク・インタフェース
２１２　第１のネットワーク・インタフェース
２１４　プロセッサ
２１６　一次記憶
２１８　二次記憶
２２０　入出力インタフェース
２２２　オペレーティング・システム
２２４　プロトコル・スタック
２２６　データグラム生成器
３０２　ノード
３０６　ルータ装置
３１０　ノード
４０２　第１のノード
４０７　Ｘ．２５パケット交換ネットワーク
４１０　ルータ
４１２　イーサネット・ローカル・エリア・ネットワーク

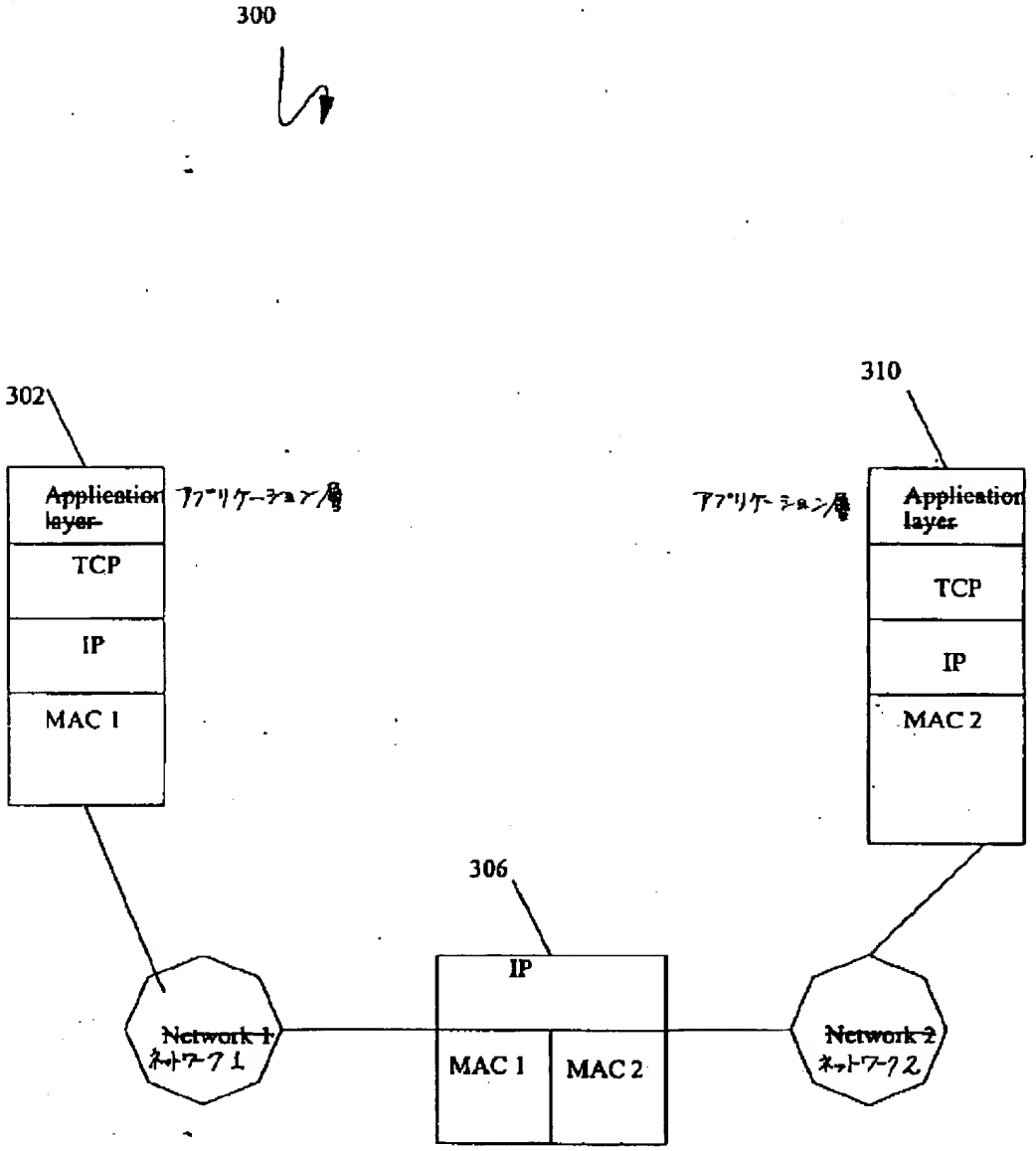
【図１】


100
102
イーサネット
Ethernet
MTU=1500
IEEE 802.3
MTU=1492
110
114
PPP
MTU=1100
送信元ノード
Source
Node
ルータ
Router
Router
ルータ
宛先ノード
Destination
Node
104
108
112
106
116
MTU=1006
Intermediate
Node
中間ノード
118

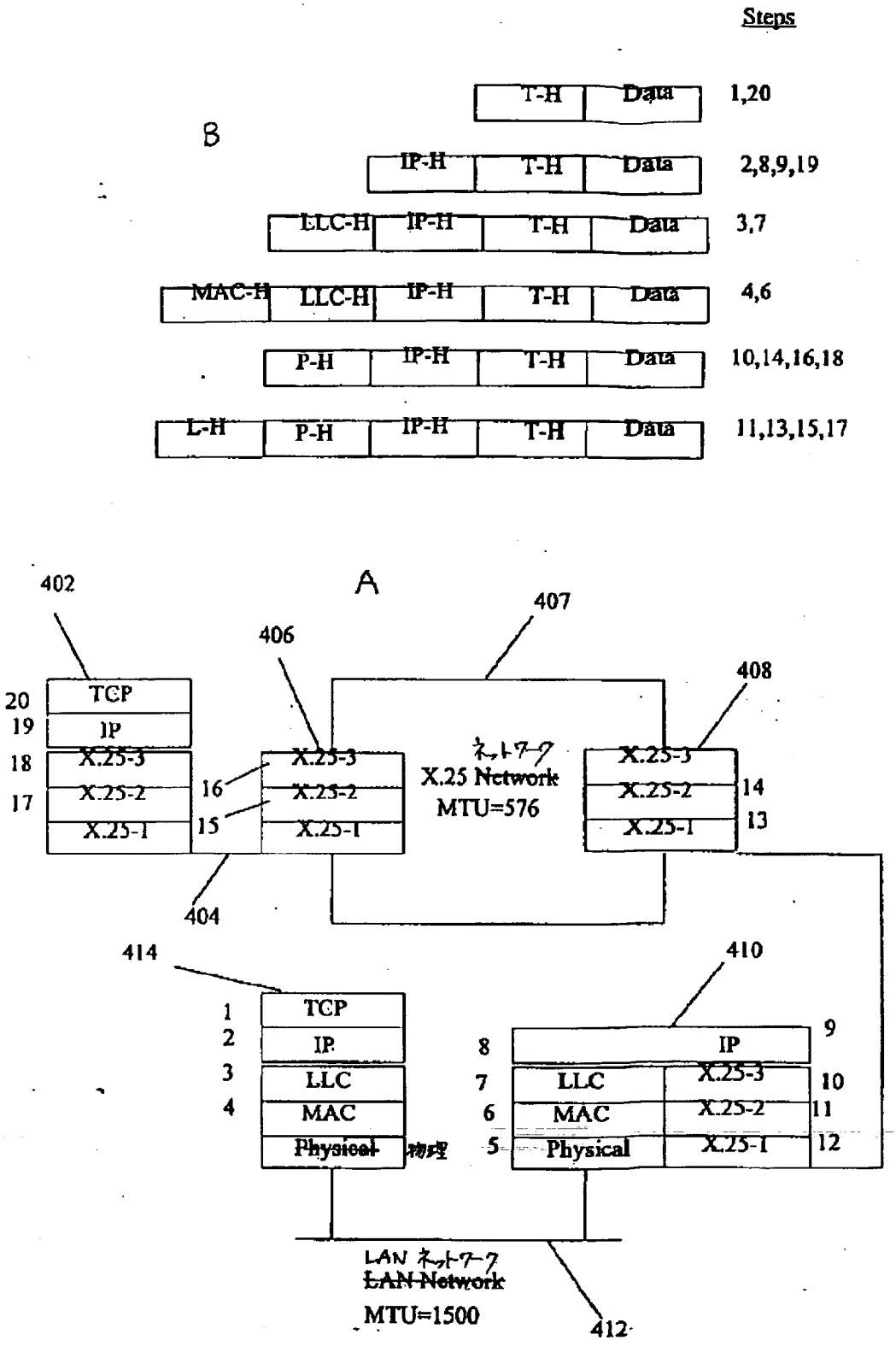
【図２】


202
216
222
218
212
211
Network Interface
ネットワーク・インターフェース
Network Interface
ネットワーク・インターフェース
Operating system
Secondary Storage
224
Protocol Stacks
226
I/O Interface
入力インタフェース
datagram generator
Primary Storage
220
Processor
214

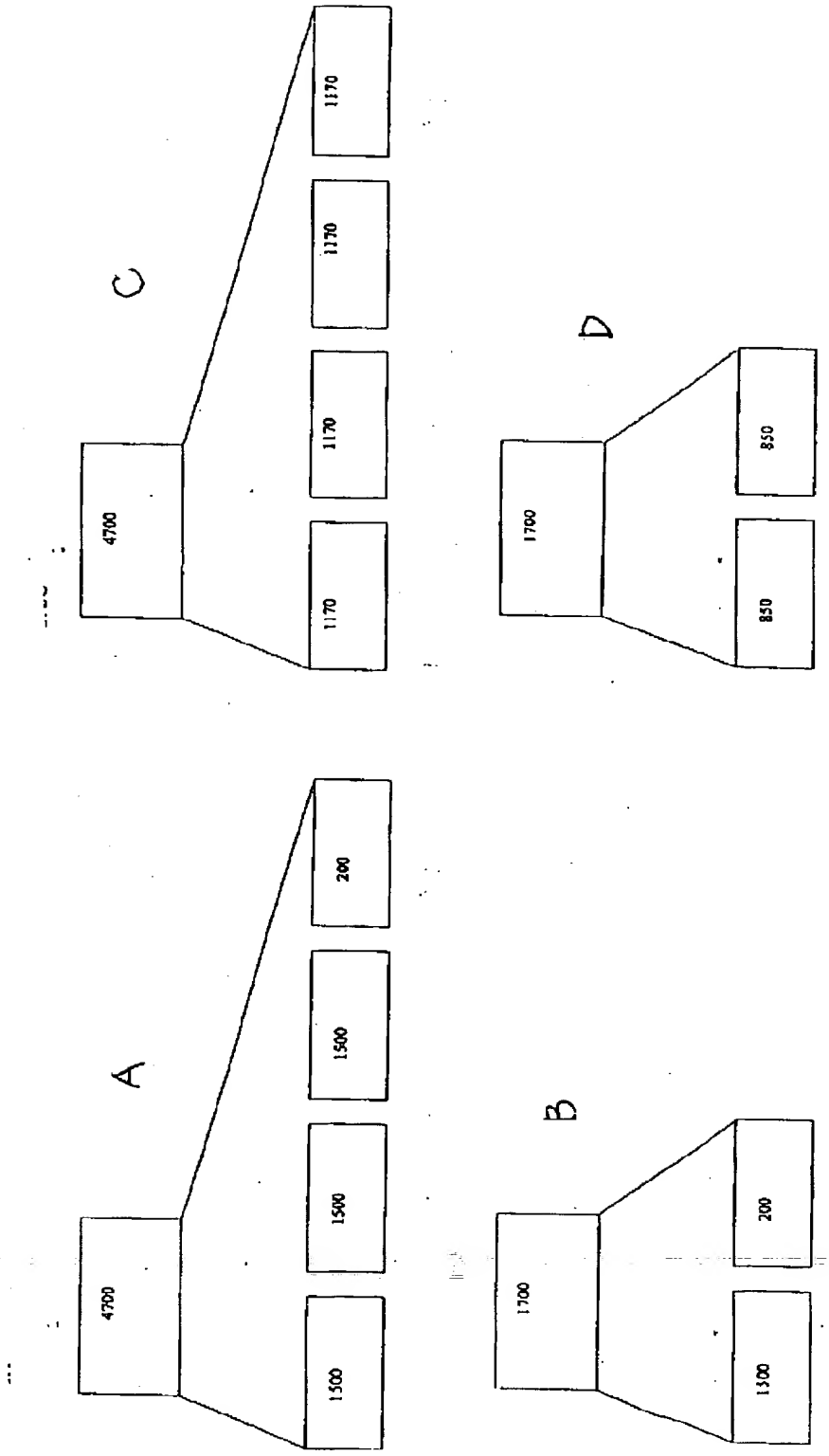
【図３】


300
302
Application layer
アプリケーション層
TCP
IP
MAC 1
310
アプリケーション層
Application layer
TCP
IP
MAC 2
306
IP
MAC 1
MAC 2
Network 1
ネットワーク1
Network 2
ネットワーク2

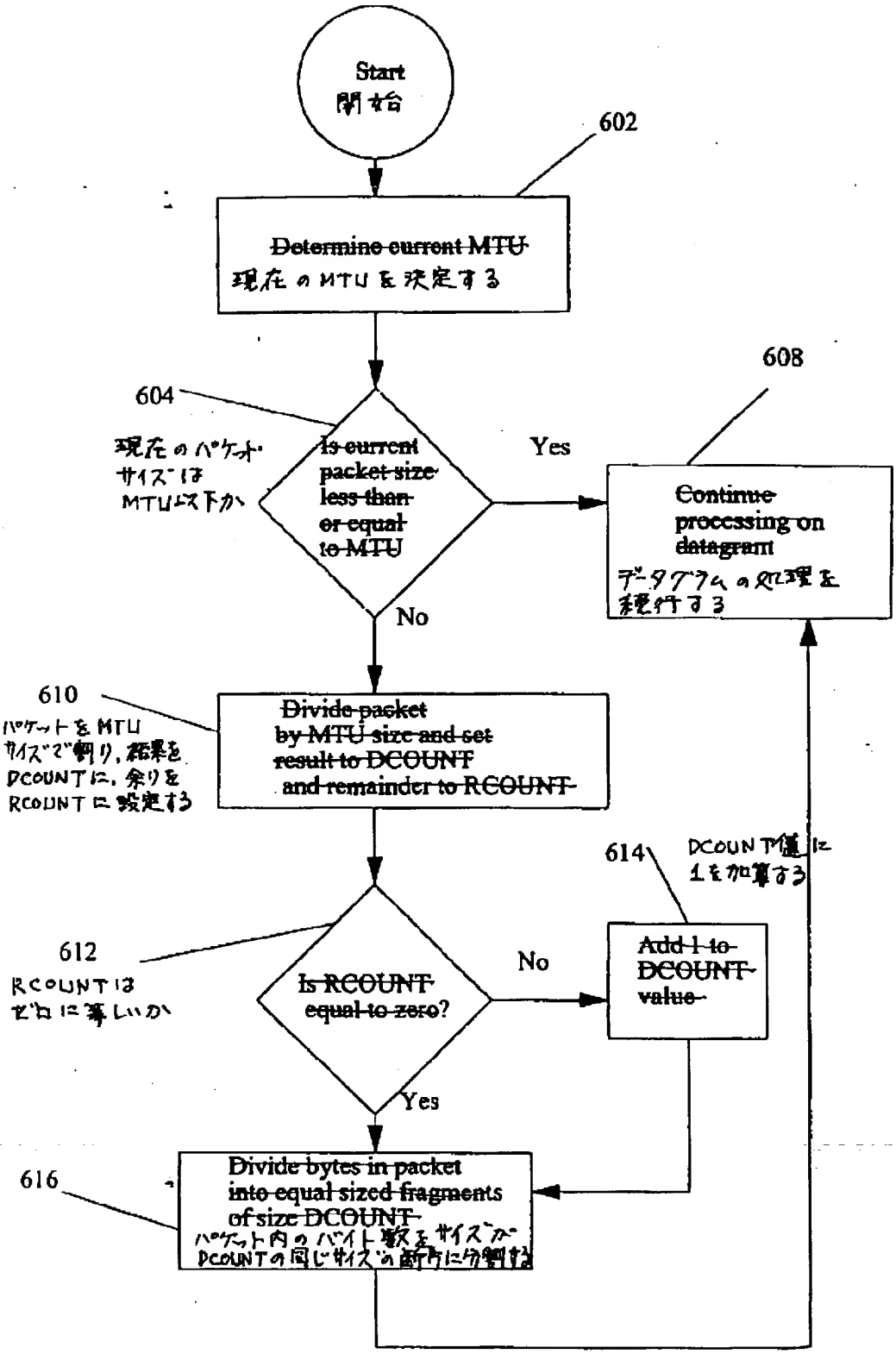
【図４】


Steps
B
T-H
Data
1,20
IP-H
2,8,9,19
LLC-H
3,7
MAC-H
4,6
P-H
10,14,16,18
L-H
11,13,15,17
A
402
406
407
408
TCP
IP
X.25-3
X.25-2
X.25-1
20
19
18
17
16
15
14
13
404
414
410
1
2
3
4
8
7
6
5
9
10
11
12
LLC
MAC
Physical
物理
X.25 Network
MTU=576
LAN Network
MTU=1500
412

【図５】

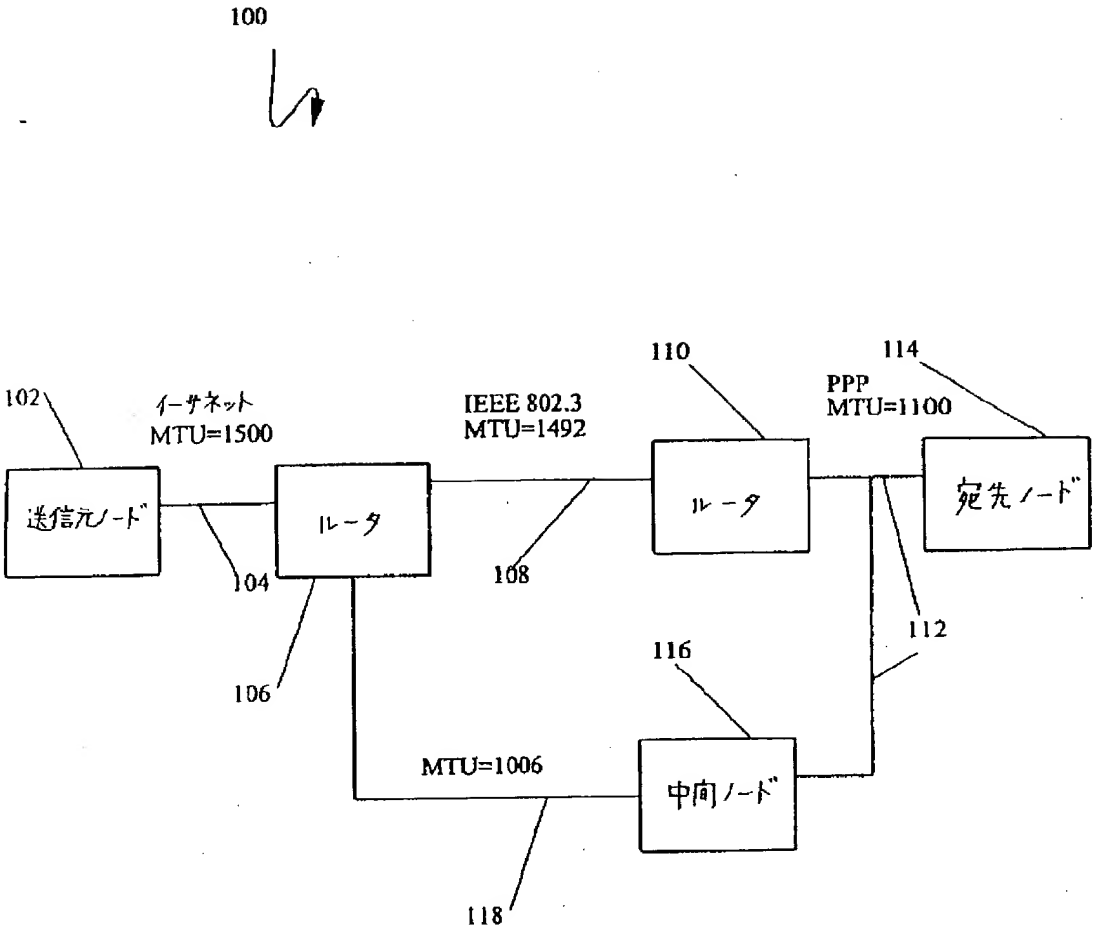

C
4700
1170
1170
1170
1170
D
1700
850
850
A
4700
1500
1500
1500
200
B
1700
1500
200

【図6】

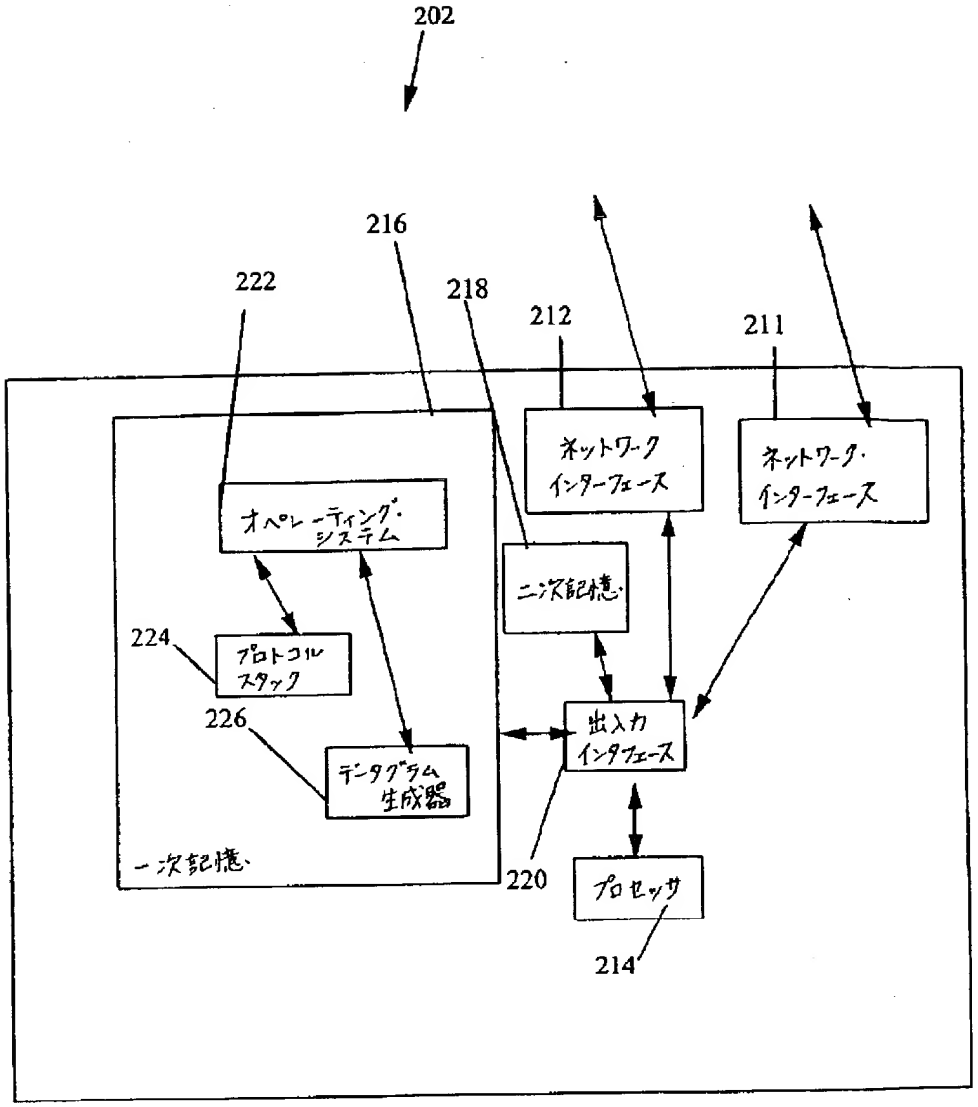

Start
開始
602
Determine current MTU
現在のMTUを決定する
604
現在のパケットサイズはMTU以下か
Is current packet size less than or equal to MTU
Yes
608
Continue processing on datagram
データグラムの処理を継続する
No
610
パケットをMTUサイズで割り、結果をDCOUNTに、余りをRCOUNTに設定する
Divide packet by MTU size and set result to DCOUNT and remainder to RCOUNT
614
DCOUNT値に1を加算する
612
RCOUNTはゼロに等しいか
Is RCOUNT equal to zero?
No
Add 1 to DCOUNT value
Yes
616
Divide bytes in packet into equal sized fragments of size DCOUNT
パケット内のバイト数をサイズがDCOUNTの同じサイズの断片に分割する

【手続補正書】
【提出日】平成１０年８月５日
【手続補正１】
【補正対象書類名】図面
【補正対象項目名】全図
【補正方法】変更
【補正内容】

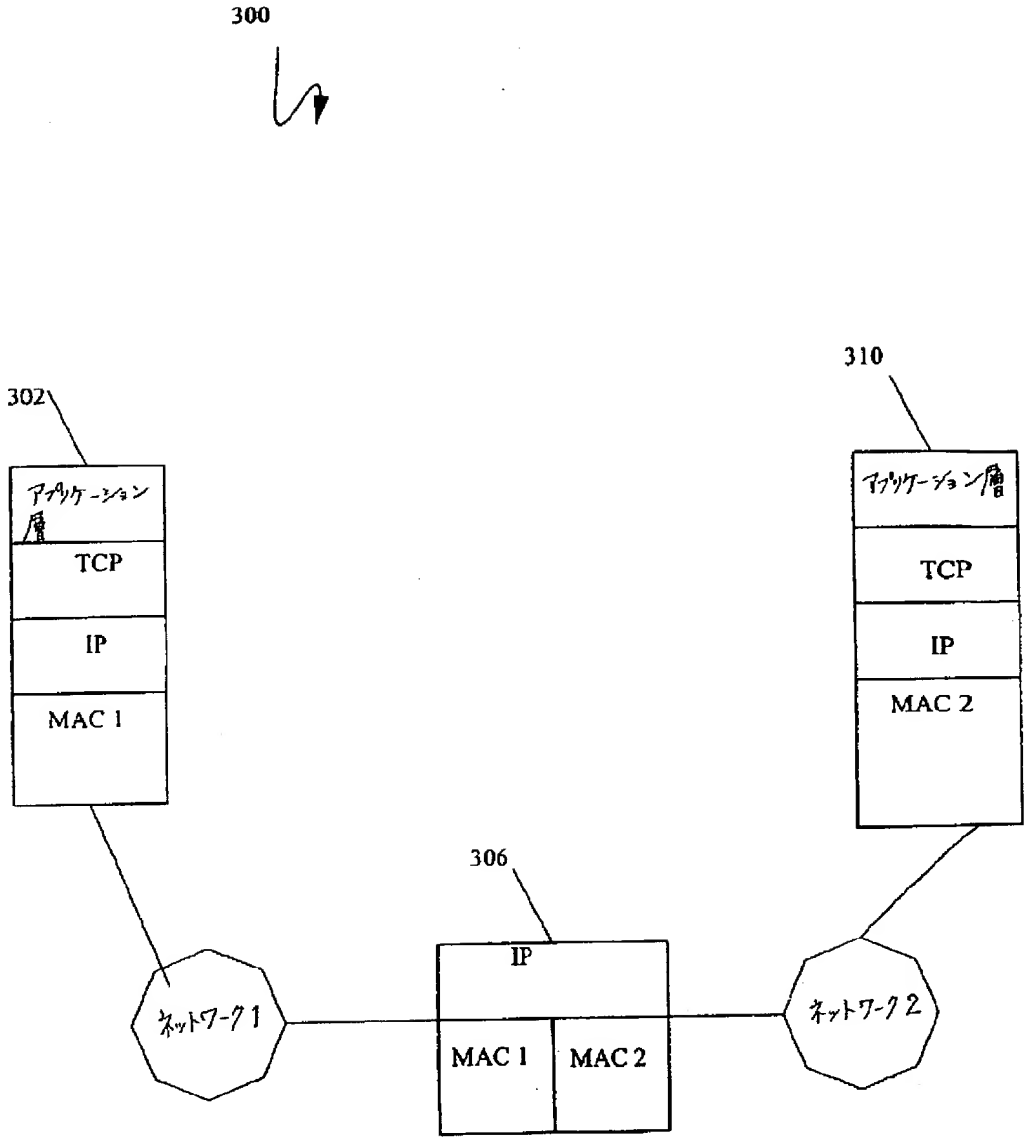
【図１】



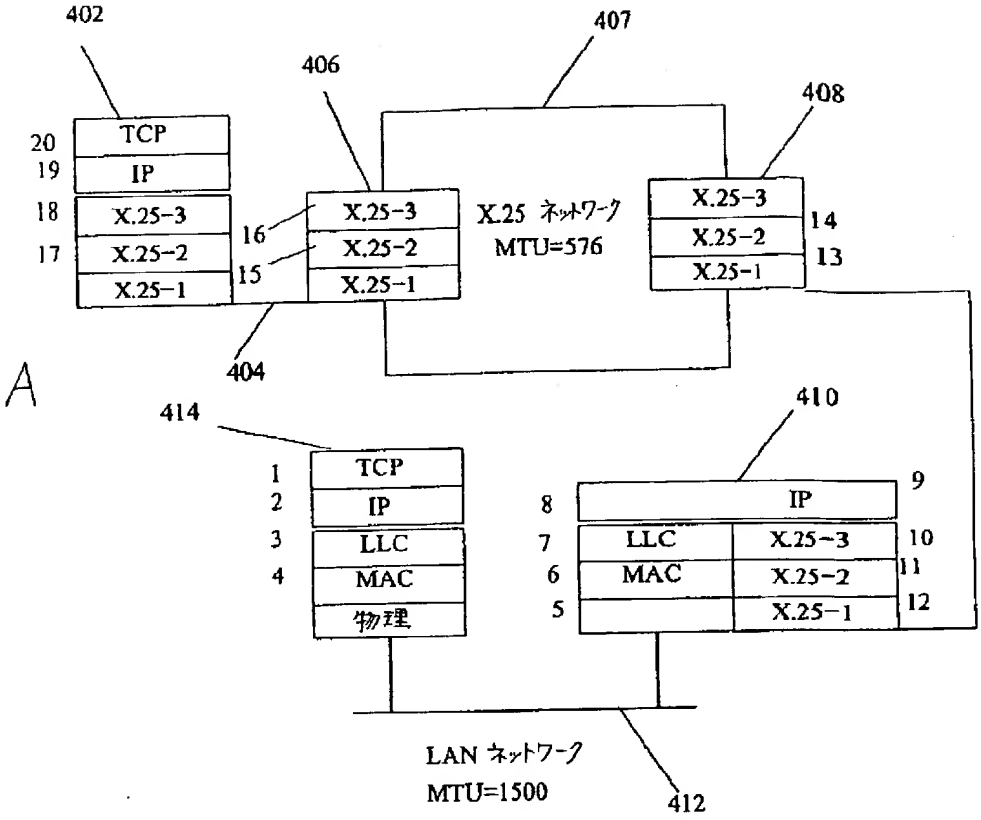
【図２】

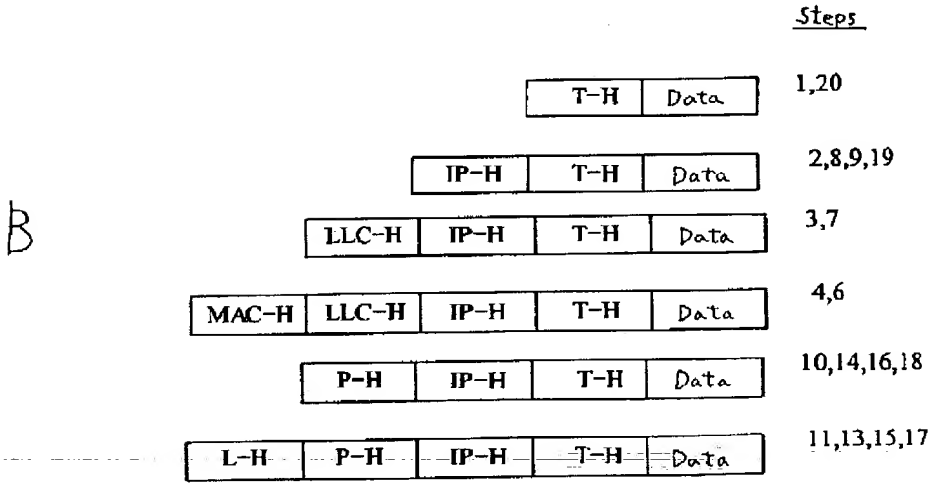

202
216
222
218
212
211
ネットワーク
インターフェース
ネットワーク
インターフェース
オペレーティング・
システム
二次記憶
224
プロトコル
スタック
226
出入力
インタフェース
テンタグラム
生成器
一次記憶
220
プロセッサ
214

【図３】

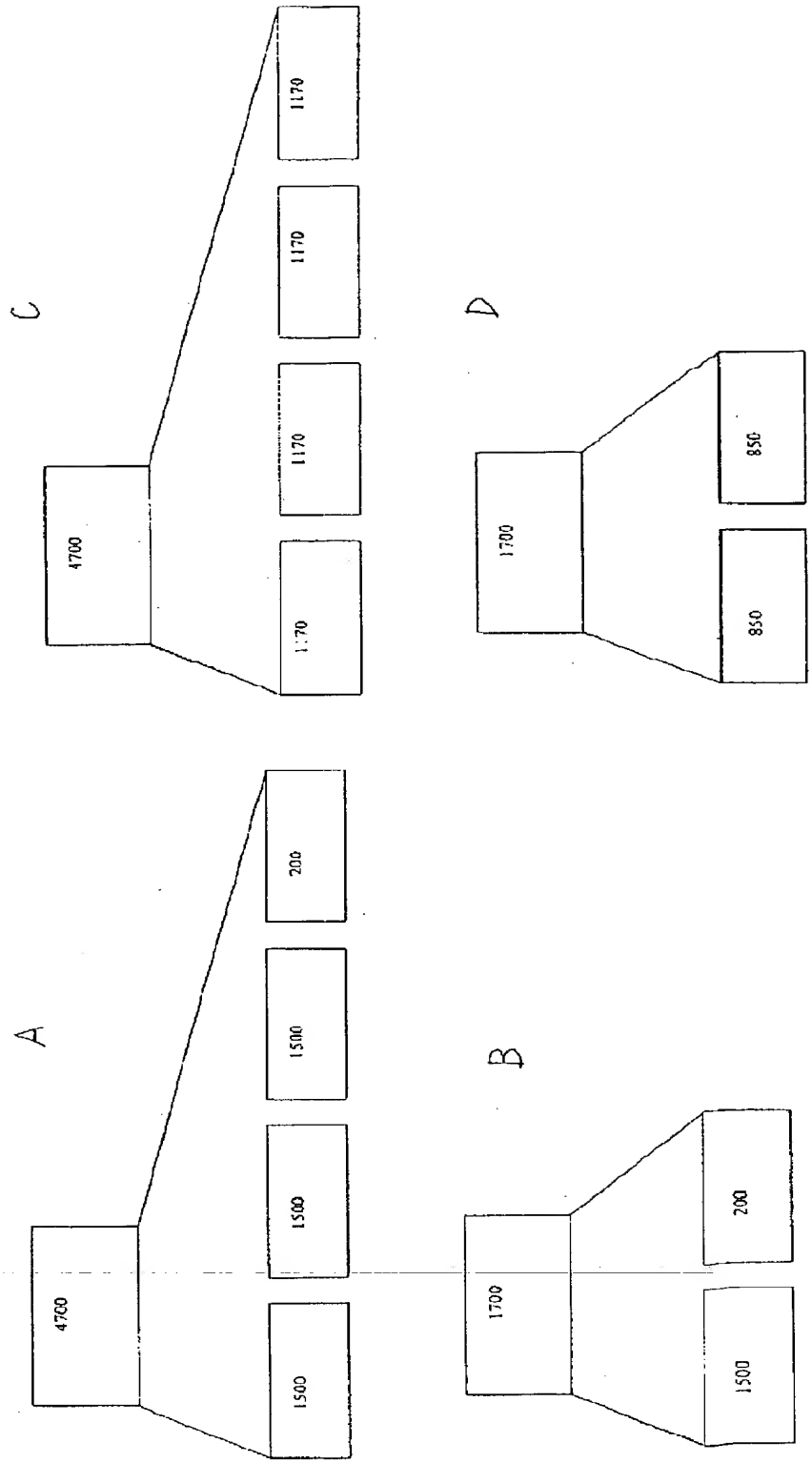

300
302
アプリケーション層
TCP
IP
MAC 1
310
アプリケーション層
TCP
IP
MAC 2
306
IP
MAC 1
MAC 2
ネットワーク1
ネットワーク2

【図４】

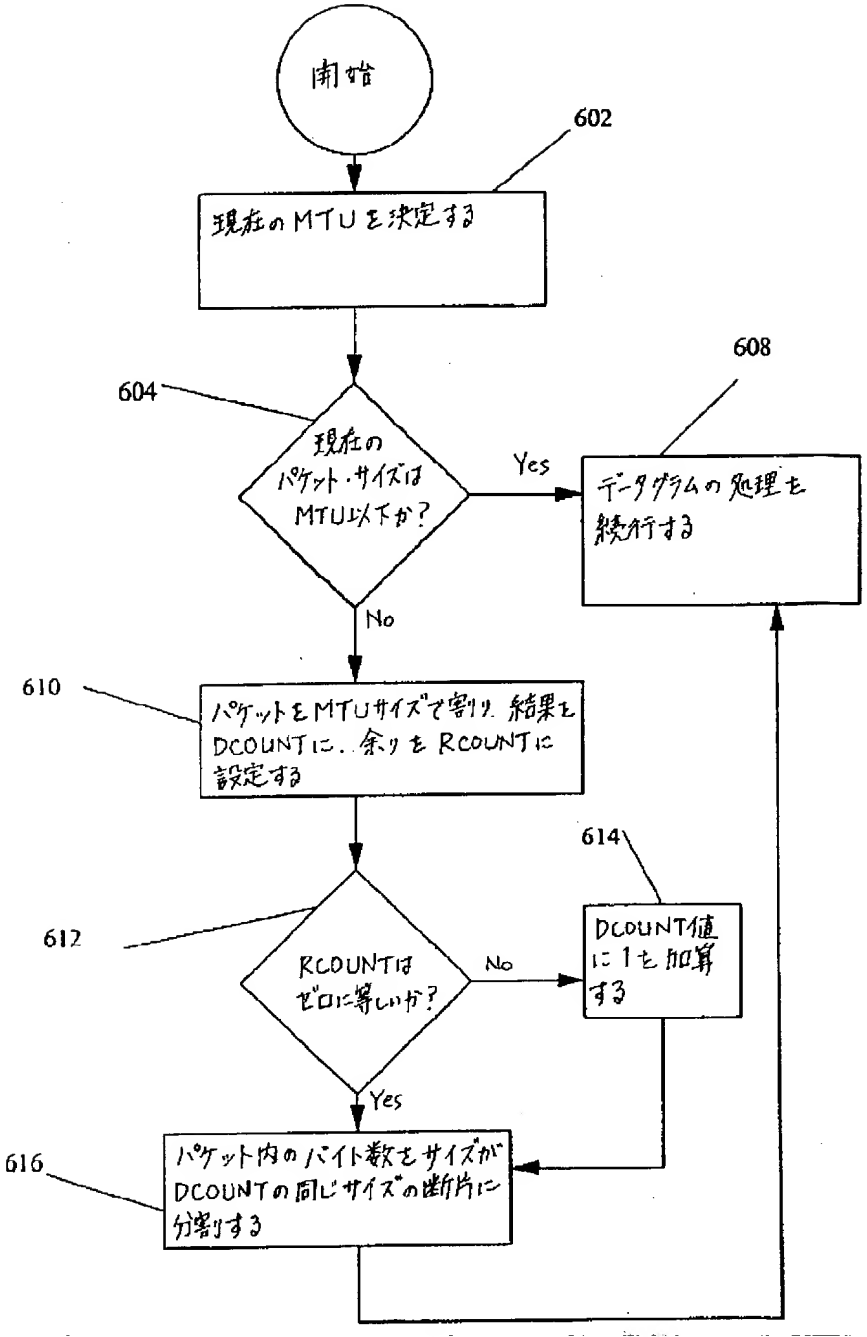

402
407
406
408
20
19
TCP
IP
18
X.25-3
17
X.25-2
X.25-1
16
15
X.25-3
X.25-2
X.25-1
X.25 ネットワーク
MTU=576
X.25-3
X.25-2
X.25-1
14
13
404
A
414
410
1
2
3
4
TCP
IP
LLC
MAC
物理
8
7
6
5
IP
LLC
MAC
X.25-3
X.25-2
X.25-1
9
10
11
12
LAN ネットワーク
MTU=1500
412


Steps
T-H
Data
1,20
IP-H
T-H
Data
2,8,9,19
B
LLC-H
IP-H
T-H
Data
3,7
MAC-H
LLC-H
IP-H
T-H
Data
4,6
P-H
IP-H
T-H
Data
10,14,16,18
L-H
P-H
IP-H
T-H
Data
11,13,15,17

【図５】


A
4700
1500
1500
1500
200
B
1700
1500
200
C
4700
1170
1170
1170
1170
D
1700
850
850

【図６】



フロントページの続き

(71)出願人 591064003
901 SAN ANTONIO ROAD
PALO ALTO, CA 94303, U.
S. A.